Electronic Supplementary Material (ESI) for CrystEngComm.

Testing the limits of NMR crystallography: the case of caffeine-citric acid hydrate – Supplementary Information

Hannah E. Kerr[a], Helen E. Mason[a], Hazel A. Sparkes[b], and Paul Hodgkinson[a]

[a] *Department of Chemistry, Durham University, Science Site, Durham DH1 3LE*
[b] *School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS*

Table S1. Time taken and number of iterations required to complete geometry optimisations (averaged over all 16 simulated-disorder structures)

			Errors		
Geometry Optimisation Method	Average CPU time / hours	Average number of iterations	Complex landscape	Bisection search failed	Repeat reset of inverse Hessian
[1]	8	49	3	0	0
[2]	48	113	7	0	1
[2-D]	24	116	11	4	5
[3]	106	373	0	0	0
[3-D]	44	126	10	4	2
[3-DO]	61	142	13	7	5
[3-DW]	67	198	10	3	14
[3-DC]	473	694	0	0	0

Table S1 shows that the calculation time and number of iterations in the optimisation increases as the convergence conditions are tightened. The calculation times are highly variable, reflecting variations in performance of the HPC platform, and so the number of iterations is the more meaningful parameter. Unsurprisingly, variable unit cell optimisations take the longest time and require the largest number of iterations to converge.

Increasing the tightness of the convergence conditions causes the number of warnings in the optimisation to increase, particularly the "complex landscape" warning. The over-tightening of the force tolerance probably contributes to the increased number of warnings in [3-D]–[3-DW]. However, no warnings were observed in [3], despite the large number of iterations required. In comparison, no dispersion correction was used in [2] and the convergence criteria were looser but "complex landscape" warnings were still present in nearly half of the optimisations.

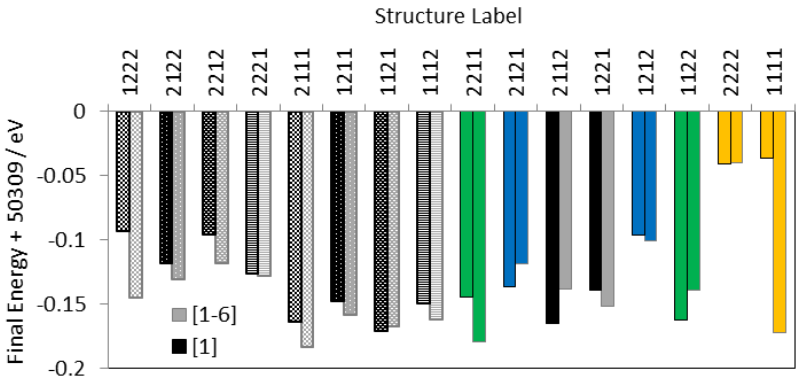

Figure S1. The energy of the simulated-disorder structures following the initial crude optimisation in CASTEP v6.0 [1-6] and optimisation [1].

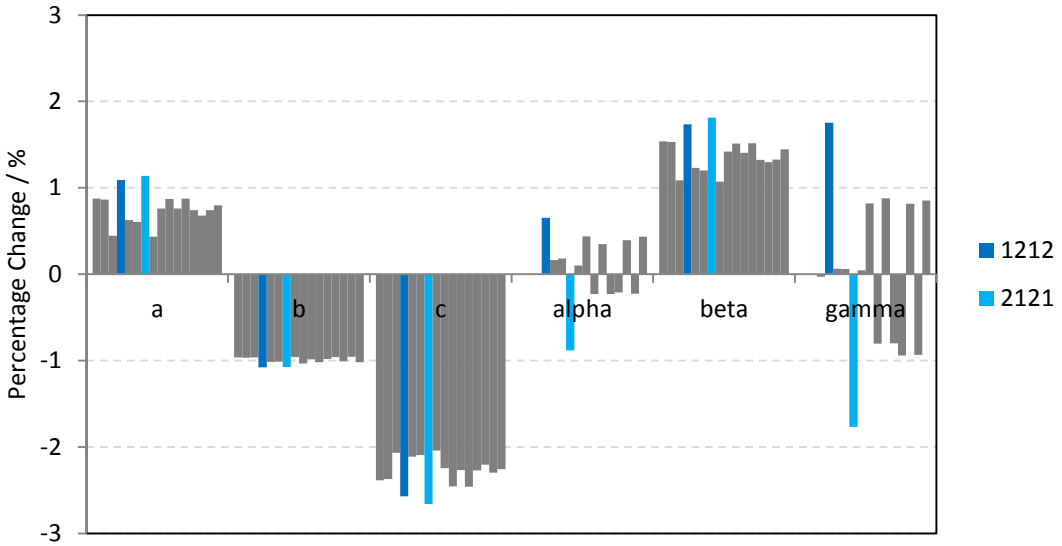

Figure S2. Percentage change of the unit cell parameters of the 16 simulated disorder structures following geometry optimisation with variable unit cell parameters (relative to unoptimised value). The 1212 symmetry-related pair is highlighted.

The variation in unit cell parameters for the 16 simulated disorder structures was relatively small, with the majority of the parameters changing by roughly 1%. The unit cell length along the *c* axis reduced by a moderate 2% in all cases during the optimisations. The **1212** pair showed a significant lowering in final energy relative to the other structures, which is reflected by the particularly large change in the angles, which all deviate away from 90˚ upon optimisation. Apart from the **1111** pair, which retains a monoclinic crystal system, the symmetry of the other structures drops to triclinic. The sign of the change in α and β angles from 90˚ is not physically significant.

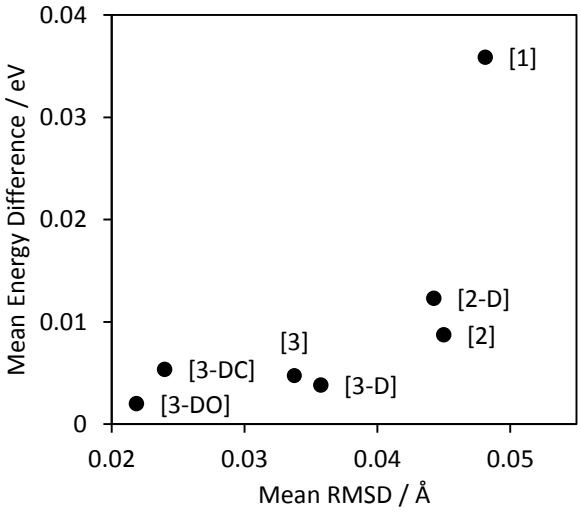

Figure S3. Correlation between the mean energy difference and mean heavy atom RMSD between symmetry-related pairs following the different geometry optimisations.

Fig. S3 shows that the total heavy atom RMSD is correlated with the mean energy difference between symmetry-related pairs as expected. [3-DC] and [3-DO] result in the lowest RMSDs between symmetry-related pairs, though [3-DC] is not considered to be as physical, as discussed in the main text.

Linewidth Simulation

The homogeneous contribution to the ^{13}C linewidth consists of a "nonrefocussable linewidth", $\Delta' = \frac{1}{\pi T_2'}$, assumed to be Lorentzian in nature, where T_2', is the time-constant for decay of the ^{13}C magnetisation under spin-echo conditions. T_2' values and subsequently Δ' values were calculated from spin-echo experiments, as described the main text, for both CCA and the 1:1 cocrystal.

Table S2. Breakdown of the linewidths and contributions for CCA. Errors on the final digits are given in brackets.

Caffeine Label	No. neighbouring N atoms	CCA Δ_{exp} / Hz	CCA Δ' / Hz
C10	1	87(1)	20(5)
C10i	1	85(1)	17(2)
C12	1	87(1)	18(3)
C9	1	91(1)	16(3)
C11	2	104(1)	17[a]
C8	2	107(1)	18(4)
C7	2	112(1)	16(2)
C8i	2	107(1)	16(5)

[a] S/N too poor for accurate determine. Average value of other Δ' values used.

The inhomogeneous contribution, $\Delta_{inhomog}$, includes the broadening due to disorder, ABMS and quadrupolar effects from the nitrogen atoms in caffeine. The latter effects were simply estimated to be the same as those of the 1:1 ordered caffeine-citric acid cocrystal, and calculated by subtracting Δ' from the experimental linewidths for the 1:1 cocrystal, see Table S3. Although the ABMS contribution will be different between the system, they are constant for a given sample and so do

not affect the quality of correlation in Fig. 6. The inhomogeneous contributions were modelled using a Gaussian lineshape.

Table S3. A breakdown of the ^{13}C linewidth contributions of the ordered 1:1 cocrystal. Uncertainties on the final digits are given in brackets.

Caffeine Label	1:1 cocrystal Δ_{exp} / Hz	1:1 cocrystal Δ' / Hz	1:1 cocrystal $\Delta_{inhomog}$ / Hz
C10	76(1)	9(1)	67(1)
C10i	80(1)	11(1)	73(1)
C12	69(1)	9(1)	60(1)
C9	85(1)	5.5(5)	80(1)
C11	104(1)	9(1)[a]	102(1)
C8	100(1)	9(1)[a]	90(1)
C7	100(1)	4.5(2)	97(1)
C8i	75(1)	4.8(6)	70(1)

[a] C8/11 for the 1:1 cocrystal overlap so the T_2' values could not be distinguished separately.

The site-specific Δ' values for CCA (Table S2) and $\Delta_{inhomog}$ values from the 1:1 cocrystal (Table S3) were used as the input linewidths for a Python script that used pNMRsim[1] to calculate synthetic ^{13}C spectra for each caffeine site from the isotropic chemical shifts read from the 16 .magres files using the MagresPython[2] library. The sum of these spectra over the 16 disorder structures gave a single spectrum per site, the linewidths of which were deconvoluted in Gsim[3] and tabulated in Table S4.

Table S4. Linewidths measured from simulated spectra from the [3-DO] optimisation summed over all 16 simulated-disorder structures.

Caffeine Label	Δ_{sim} / Hz
C10	84.6
C10i	82.9
C12	76.1
C9	91.7
C11	119.5
C8i	106.2
C7	106.9
C8	87.3

(1) P. Hodgkinson, *pNMRsim: a general simulation program for large problems in solid-state NMR*, www.dur.ac.uk/paul.hodgkinson/pNMRsim, accessed 2016.
(2) S. Sturniolo, T. F. G. Green, R. M. Hanson, M. Zilka, K. Refson, P. Hodgkinson, S. P. Brown and J. R. Yates, *Solid State Nucl. Magn. Reson.*, **2016**, 10.1016/j.ssnmr.2016.05.004.
(3) V. Zorin, *GSim - a visualisation and processing program for solid-state NMR*, www.sourceforge.net/projects/gsim/, accessed 2016.